

University of Dundee

## Application of pharmacogenomics and bioinformatics to exemplify the utility of human ex vivo organoculture models in the field of precision medicine

Cowan, Karen; MacLuskie, Graeme; Finch, Michael; Palmer, Colin N. A.; Hair, Jane; Bylesjo, Max

*Published in:*  
PLoS ONE

*DOI:*  
[10.1371/journal.pone.0226564](https://doi.org/10.1371/journal.pone.0226564)

*Publication date:*  
2019

*Licence:*  
CC BY

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

### *Citation for published version (APA):*

Cowan, K., MacLuskie, G., Finch, M., Palmer, C. N. A., Hair, J., Bylesjo, M., Lynagh, S., Brankin, P., McNeil, M., Low, C., Mallinson, D., Gourlay, E. M., Child, H., Cheyne, L., & Bunton, D. C. (2019). Application of pharmacogenomics and bioinformatics to exemplify the utility of human ex vivo organoculture models in the field of precision medicine. *PLoS ONE*, 14(12), 1-15. [e0226564]. <https://doi.org/10.1371/journal.pone.0226564>

### **General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

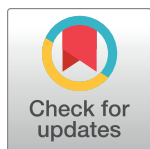
RESEARCH ARTICLE

# Application of pharmacogenomics and bioinformatics to exemplify the utility of human *ex vivo* organoculture models in the field of precision medicine

Karen Cowan<sup>1\*</sup>, Graeme Macluskie<sup>1</sup>, Michael Finch<sup>1</sup>, Colin N. A. Palmer<sup>2</sup>, Jane Hair<sup>3</sup>, Max Bylesjo<sup>4</sup>, Sarah Lynagh<sup>4</sup>, Pamela Brankin<sup>5</sup>, Marian McNeil<sup>6</sup>, Carolyn Low<sup>6</sup>, David Mallinson<sup>7</sup>, Elaine M. Gourlay<sup>7</sup>, Hannah Child<sup>6</sup>, Linda Cheyne<sup>6</sup>, David C. Bunton<sup>1</sup>

**1** REPROCELL Europe Ltd, Thomson Pavilion, Glasgow, Scotland, United Kingdom, **2** School of Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland, United Kingdom, **3** NHS Greater Glasgow & Clyde, Queen Elizabeth University Hospital, Glasgow, Scotland, United Kingdom, **4** Fios Genomics Ltd, Nine Edinburgh Bioquarter, Edinburgh, Scotland, United Kingdom, **5** Aridhia Informatics Ltd Teaching and Learning Building, Queen Elizabeth University Hospital, Glasgow, Scotland, United Kingdom, **6** Stratified Medicines Scotland Innovation Centre, Teaching and Learning Building, Queen Elizabeth University Hospital, Glasgow, Scotland, United Kingdom, **7** Sitemic Ltd, West of Scotland Science Park, Glasgow, Scotland, United Kingdom

\* [karen.cowan@reprocell.com](mailto:karen.cowan@reprocell.com)



## OPEN ACCESS

**Citation:** Cowan K, Macluskie G, Finch M, Palmer CNA, Hair J, Bylesjo M, et al. (2019) Application of pharmacogenomics and bioinformatics to exemplify the utility of human *ex vivo* organoculture models in the field of precision medicine. PLoS ONE 14(12): e0226564. <https://doi.org/10.1371/journal.pone.0226564>

**Editor:** Erika Cecchin, Centro di Riferimento Oncologico, ITALY

**Received:** May 8, 2019

**Accepted:** November 28, 2019

**Published:** December 20, 2019

**Copyright:** © 2019 Cowan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data can be found in the Figshare repository and has the following reserved DOI: [10.6084/m9.figshare.10101371](https://doi.org/10.6084/m9.figshare.10101371).

**Funding:** Funding for this study was via a project grant from Stratified Medicines Scotland Innovation Centre (SMS-IC, <https://www.stratmed.co.uk/>), provided to REPROCELL Europe Ltd (<https://www.reprocell.com/>) and the University of Dundee (<https://www.dundee.ac.uk/>). REPROCELL

## Abstract

Here we describe a collaboration between industry, the National Health Service (NHS) and academia that sought to demonstrate how early understanding of both pharmacology and genomics can improve strategies for the development of precision medicines. Diseased tissue ethically acquired from patients suffering from chronic obstructive pulmonary disease (COPD), was used to investigate inter-patient variability in drug efficacy using *ex vivo* organocultures of fresh lung tissue as the test system. The reduction in inflammatory cytokines in the presence of various test drugs was used as the measure of drug efficacy and the individual patient responses were then matched against genotype and microRNA profiles in an attempt to identify unique predictors of drug responsiveness. Our findings suggest that genetic variation in CYP2E1 and SMAD3 genes may partly explain the observed variation in drug response.

## Introduction

It is well recognised that one size does not fit all when it comes to the treatment of many diseases. Getting the right drug to the right patient at the right dose has become the focus of precision medicine, which provides hope that patients may receive the most appropriate treatment sooner, improving their quality of life and reducing the support required from health care systems and wider society [1]. Health economists are recognising the potential of precision medicine and are beginning to apply the concept to their research [2].

Europe, SMS-IC and Aridhia (<https://www.aridhia.com/>) provided in-kind contributions to research materials and staff costs (REPROCELL: KC, GM, MF and DCB; SMS IC: MM, CL, HC, LC; Aridhia: PB). Fios Genomics (<https://www.fiosgenomics.com/>) was sub-contracted by REPROCELL to conduct bioinformatics analysis. Sismic (<http://www.sismic.co.uk/>) was sub-contracted by REPROCELL to conduct RNA, DNA extractions and miRNA analysis. SMS-IC is funded by the Scottish Funding Council. The funding organisations did not play an additional role in the study design, data collection and analysis, decision to publish or preparation of the manuscript and only provided financial support in the form of authors' salaries and research materials. The specific roles of the authors are articulated in the 'authors contributions' section.

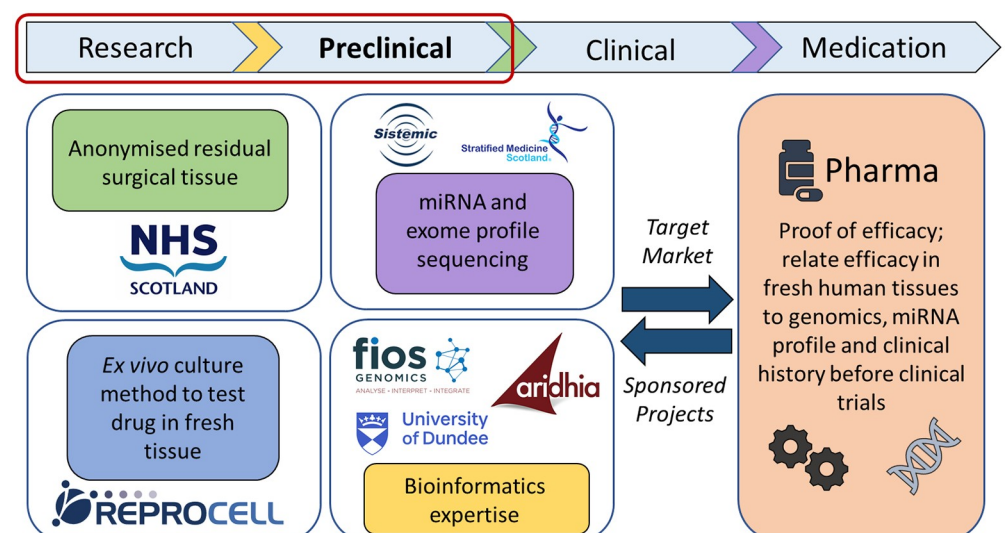
**Competing interests:** REPROCELL Europe Ltd (<https://www.reprocell.com/>) is a commercial provider of laboratory based tests for preclinical research. KC, GM, MF and DCB are all paid employees of REPROCELL Europe. Aridhia (<https://www.aridhia.com/>) is a commercial provider of digital research environments for industry and academia. PB is a paid employee of Aridhia. Fios Genomics (<https://www.fiosgenomics.com/>) is a commercial provider of bioinformatics services. MB and SL are both paid employees of Fios Genomics. Sismic (<http://www.sismic.co.uk/>) is a commercial provider of miRNA focussed services. DM and EMG are both paid employees of Sismic. These commercial affiliations do not alter adherence of the authors to PLOS ONE policies on sharing data and materials.

The genomics revolution has underpinned much of this research. As the cost of gene sequencing has fallen, the ability to rapidly identify an individual's genotype as part of routine health care has become possible. For precision medicines to be developed however, genomics must be linked to pharmacology, thereby relating the individuals genotype to the effectiveness, potency and tolerability of a drug. It is through pharmacogenomics that truly personalised therapies may emerge, yet the link between genomics and pharmacology may not be properly understood until expensive and risky clinical trials are conducted.

Here we describe a collaboration between industry, the National Health Service (NHS) and academia that sought to demonstrate how an early understanding of both pharmacology and genomics can improve strategies for the development of precision medicines. By using the latest pharmacology techniques in human fresh tissues donated by the target patient population, combined with genomics and clinical metadata associated with each individual, an improved understanding of the link between genetics and inter-individual drug responses emerges. A diagrammatic representation of this collaboration has been provided (Fig 1) to detail the interaction between each collaborator and the different data streams.

An early understanding of patient stratification during drug discovery is becoming increasingly important. Selection and optimisation of candidate drugs for well-defined patient subsets has the potential to help in the design of more rapid, targeted clinical trials.

A key incentive to better understand pharmacogenomics during the drug discovery process is the rapid increase in drug development costs. The most recent estimates of the out-of-pocket costs (i.e. excluding capital costs) of drug development are in the region of \$890 million [3], with approximately 70% of the costs incurred during clinical development. The most common cause of failure is poor efficacy at phase II or III [3–6], which is in part attributed to 'all comers' clinical trials of entire patient populations that are likely to include both 'responders' and 'non-responders' to the drug being tested. Precision medicine aims to improve the prediction of clinical efficacy by selecting for clinical trials only those patient sub-populations likely to gain clear benefit. This approach also benefits patients classed as 'non-responders' for a drug treatment as it opens up clear market opportunities for drug developers to develop effective treatments for these populations. Predictions of optimal patient sub-populations are however dependent on the quality of the information available to stratify these sub-populations before a



**Fig 1. Diagram describing the interaction between each collaborator and the different data streams.**

<https://doi.org/10.1371/journal.pone.0226564.g001>

drug enters clinical development. Preclinical tests of drug effects must therefore closely reflect the intended patient population(s).

The most desired traits in preclinical models are “physiological relevance” and the ability to translate findings to likely clinical responses [4,6–8]. Obtaining a high level of confidence in the biological role of the drug target in human disease during preclinical development has been shown to be a good predictor of successful clinical trials [6]. Human fresh tissues and complex human 3D tissue models that reflect the biology of disease are therefore increasingly being used by Pharma to improve the prediction of efficacy in clinical trials [7,9,10]. In this study human lung tissue from diseased patients was used to further increase confidence in the ‘physiological relevance’ of the test system. Using human tissue based systems at a preclinical stage also provides the opportunity to allow the discovery of biomarkers that can be leveraged in the clinic, either as a marker of drug response or as a patient stratification marker. Although the data between different patients can be variable, this is viewed as an opportunity for an early understanding of the extent and causes of inter-patient variation in drug response.

Chronic Obstructive Pulmonary Disease (COPD) is a major health problem and is an example of a complex condition, with many clinical phenotypes. Many patients receive minimal clinical benefit from common medications, most likely due to the combination of variations in disease subtype and genotype.

In this project, diseased tissue ethically acquired from patients suffering from COPD, was used to investigate inter-patient variability in drug efficacy using *ex vivo* organocultures of fresh lung tissue as the test system. To investigate the level of patient stratification that can be achieved from a preclinical study, the commonly used ‘all comers’ clinical trial patient recruitment policy [11] was mimicked with no donor inclusion criteria other than a clinical diagnosis of COPD applied.

In order to assess patient variation in responsiveness to both ‘standard of care’ and combination ‘standard of care’ therapies, the reduction in inflammatory cytokines in the presence of the test drugs was used as the measure of drug efficacy.

Included in this study were 3 different ‘standard of care’ medications that reflect what is currently available to patients and what is routinely prescribed to those suffering with differing severities and phenotypes of COPD. Roflumilast is a selective, long-acting inhibitor of the enzyme phosphodiesterase-4 (PDE-4) commonly used to treat COPD patients that present with chronic cough. It has been shown to provide clinical benefit however has been associated with multiple adverse events [12]. Fluticasone is commonly an inhaled corticosteroid used in the treatment of asthma and COPD, however variation in treatment efficacy has been reported [13]. Formoterol is a long-acting  $\beta_2$  adrenoceptor agonist, intended for use as a bronchodilator in the management of COPD. The individual patient *ex vivo* lung tissue responses were then matched against genotype, patient demographic and microRNA profiles to identify unique predictors of drug responsiveness and demonstrate the combined power of pharmacology and genomics during preclinical development.

## Materials and methods

### *Ex vivo* organoculture—REPROCELL

Lung parenchyma tissue was ethically obtained from 25 patients, clinically diagnosed with COPD who were undergoing therapeutic resection surgery for cancer or COPD. Residual tissue from the resection surgery, which was not required for clinical pathology, was acquired from NHS Research Scotland Biorepository Network and also through the REPROCELL tissue network. The West of Scotland Research Ethics Committee (12/ws/0069) granted approval for this study, and patients provided written consent complying with the declaration of Helsinki.

Two biopsies (5 mm<sup>3</sup>) were immediately processed in *RNAlater*, prior to storage and subsequent RNA and DNA extraction. The remaining biopsies (5 mm<sup>3</sup>) were subjected to the following culture protocol.

Culture medium was prepared by supplementing RPMI 1640 culture medium with the following constituents: gentamicin (100 µg/mL), amphotericin B (0.625 µg/mL), FBS (0.5%), retinyl acetate (0.1 µg/mL), bovine insulin (1 µg/mL) and nystatin (1 µg/mL).

Following an overnight rest period in culture medium, lung biopsies (2 per well) were cultured in a humidified incubator (37 °C, 5% CO<sub>2</sub>), in culture media fortified with either roflumilast (100 nM), fluticasone (1 µM), formoterol (10 nM), roflumilast (100 nM) + fluticasone (1 µM), formoterol (10 nM) + roflumilast (100 nM) or the DMSO vehicle control. LPS (100 ng/mL) was added to all wells, to mimic the effects of a bacterial exacerbation of COPD and therefore boost biopsy inflammatory cytokine release. Each experimental condition was performed in duplicate culture wells.

At 24 hours post-incubation start, the supernatant from each well was collected, protease inhibitor added, and the samples stored at -80 °C prior to analysis for TNFα content.

Levels of TNFα (pg/mL) were measured in culture supernatants using a magnetic bead-based assay for the Luminex MAGPIX platform. Each culture supernatant sample was analysed in duplicate and the mean value used in downstream analyses.

### RNA, DNA extraction—Sistemic

DNA was extracted from approximately 10 mg tissue using the PureLink™ Genomic DNA Mini Kit. DNA quality control was performed using the Agilent 2200 TapeStation and the Genomic DNA ScreenTape kit to determine the DNA integrity number (DIN).

RNA was extracted from approximately 10 mg of tissue. Tissue was homogenised and total RNA was then extracted using the miRCURY RNA Isolation Kit—Cell & Plant. Absorbance ratios at 260/280 nM and 260/230 nM were determined as indicators of sample yield and purity. Further RNA quality control was performed using the Agilent 2200 TapeStation and the ScreenTape R6K kit to determine the RNA integrity number (RIN).

### miRNA analysis—Sistemic

MicroRNA (miRNA) expression levels were measured using the Agilent miRNA platform, specifically; Agilent's SurePrint G3 Human v16 microRNA 8x60K microarray slides, miRBase version 16.0.

### Exome sequencing—SMS-IC

Targeted next generation sequencing libraries were prepared using the Ion Ampliseq™ Exome RDY Kit and DNA isolated from baseline lung biopsies. Multiplexed PCR was performed to produce barcoded libraries, using 100 ng of input DNA per sample and 10 amplification cycles. The Ion AmpliSeq™ Library Kit Plus and IonXpress™ Barcode Adapters were used in library preparation, according to the manufacturer's instructions. Final library concentrations were determined by quantitative real time PCR using the Ion Library TaqMan™ Quantitation Kit. Libraries were diluted to 100 pM, and 2 libraries were subsequently pooled in equal amounts for templating on the Ion OneTouch™ 2 System, using the Ion PI™ Hi-Q™ OT2 200 kit. The Ion Proton™ NGS platform was used for sequencing of multiplexed templated libraries, using the Ion PI™ Hi-Q™ Sequencing 200 Kit and the Ion PI™ Chip Kit v3, according to the manufacturer's instructions.



### Raw data storage—Aridhia

Raw data (*ex vivo* organoculture TNF $\alpha$  levels, miRNA expression profiles and exome sequencing data) was uploaded to a secure workspace (AnalytiXagility) in Aridhia's digital research platform.

Anonymised, patient demographic data obtained from NHS Research Scotland Biorepository Network or the REPROCELL tissue network was also uploaded to the collaboration's AnalytiXagility workspace. Data could then be accessed and analysed in a secure manner by authorised users.

Fios Genomics accessed data held in the AnalytiXagility research workspace to provide bioinformatic analyses. Each dataset was analysed individually and combined to determine any significant correlations between patient demographic data, genetic polymorphisms and/or miRNA profiles and the observed organoculture assay response.

All the data sets that were used in the bioinformatic analyses have been fully anonymised and may be accessed via the following [10.6084/m9.figshare.10101371](https://doi.org/10.6084/m9.figshare.10101371).

### Organoculture bioinformatic analysis—Fios Genomics

TNF $\alpha$  levels determined for each patient sample in the organoculture assay, were subjected to quality control metrics from the ArrayQualityMetrics package in Bioconductor [14]. Assays were scored on the basis of the following parameters: maplet; boxplot and heatmap. An individual patient sample was classified as an outlier if two or more of the parameters were not met.

TNF $\alpha$  levels (pg/mL) for each experimental condition were normalised using log<sub>2</sub> ratios against the DMSO vehicle control. Relative levels of TNF $\alpha$  were then visualised using bar charts, density plots and correlation plots within R software. The aim was to identify subgroups of patients that displayed a good response to one or more of the organoculture test conditions and subgroups of patients that displayed a poor response. Response to a test condition was defined as a reduction in TNF $\alpha$  levels compared to the DMSO control test condition.

Patients were then categorised as being a high responder to treatment or a low responder to treatment for use in subsequent bioinformatic analyses.

Defined patient demographic parameters and organoculture assay response were assessed using pair-wise univariate associations between all combinations of defined parameters. Associations between categorical parameters were assessed using a chi-squared test; associations between one categorical and one continuous parameter were assessed using analysis of variance (ANOVA); associations between two continuous parameters were assessed using a Spearman correlation test.

### Exome sequence bioinformatic analysis—Fios Genomics

Torrent Mapping Alignment Program was used to provide IonTorrent AmpliSeq exome sequencing data for each patient. Data was provided as a BAM file aligned to genome reference GRCh37. Genotypes called with Torrent Variant Caller were provided as per sample VCF files.

Single nucleotide polymorphisms (SNPs) from the VCF files were merged into a multi-sample VCF and BAM files were used to set missing genotypes to homozygous reference if the read-depth of the SNP in a particular sample was less than 30. VCF files were then filtered to remove low quality SNPs.

Exploratory analysis was first performed by producing principal component analysis plots, using the SNPRelate R software package. Hierarchical clustering of the data measured dissimilarity between patient exome data [15].

The genotype for all SNPs identified from the VCF file was tested for association with the organoculture assay response, this was performed using fisher-exact tests of association within the PLINK analysis toolkit [16]. Identified SNPs included those that were known to be related to genes of interest and also novel, undescribed SNPs.

Genes of interest were identified due to a literature association with the pathology of COPD and/or as being associated with lung metabolism and/or genes that may be associated with clinical response to standard of care treatments.

Identified SNPs were also cross-referenced with SNPs listed in the Genome Wide Association Studies (GWAS) Catalog to determine if any SNP had been previously reported in a human GWAS study and, if so, it was determined if the reported association was relevant to this study.

### miRNA bioinformatic analysis—Fios Genomics

Quality control was assessed using the quality control metrics from the ArrayQualityMetrics package in Bioconductor [14] as for the organoculture assay data above.

Confounding associations between defined patient demographic parameters and miRNA expression array data were assessed using pair-wise univariate associations between all combinations of defined parameters. Associations between categorical parameters were assessed using a chi-squared test; associations between one categorical and one continuous parameter were assessed using analysis of variance (ANOVA); associations between two continuous parameters were assessed using a Spearman correlation test.

Data was normalised using quantile normalisation which produces expression measures in a log base 2 format. Array batch effects due to processing of microarray data in two separate batches were corrected using the ComBat method [17].

Statistical comparisons were performed to determine if specific miRNAs were associated with organoculture assay response: the null hypothesis being that no specific differences in miRNA expression could be detected in patients that responded well in the organoculture assay compared with patients that did not respond. Linear modelling, empirical Bayesian analysis and p-value adjustment for multiple testing (Benjamini-Hochberg) was performed using the Bioconductor Limma software package [14].

miRNAs were annotated based on their experimentally verified target genes from miRTarBase [18]. miRNAs that displayed significant differential expression (uncorrected  $p < 0.05$ ), were analysed for enrichment of target gene KEGG (Kyoto Encyclopaedia of Genes and Genomes) pathway membership using a hypergeometric test. Upregulation and downregulation of genes were analysed separately.

In the same way, miRNA target genes were analysed for enrichment of gene ontology terms.

### Bioinformatic analyses integration—Fios Genomics & University of Dundee

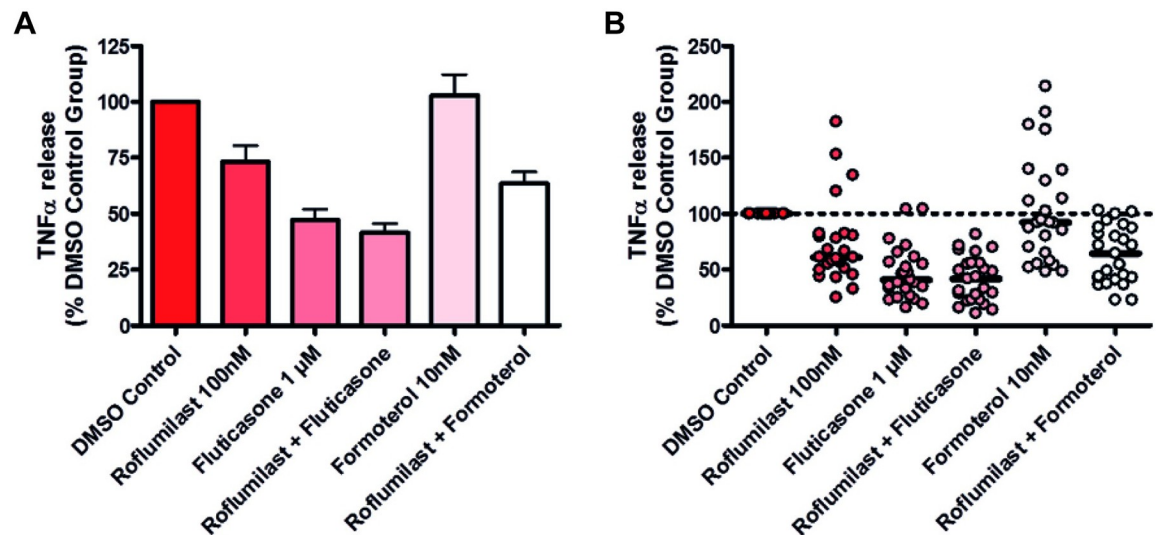
Integration of patient demographic, TNF $\alpha$  organoculture response, exome sequence and miRNA expression data.

Congruence analysis was performed by evaluating the level of overlap between all data sets. Calculations of significant overlaps were based on a hypergeometric test.

## Results

### *Ex vivo* organoculture

All patient samples passed quality control analysis as described previously. The majority of COPD patient lung samples responded to treatment with fluticasone, roflumilast or



**Fig 2. Graphs showing the effects of test articles on TNF $\alpha$  release from stimulated human lung parenchyma biopsies.**  $n = 25$  donors, all diagnosed with COPD. Data is displayed as a percentage of the corresponding DMSO control group in both graphs. A: Bar graph depicting mean + SEM TNF $\alpha$  release. B: Scatter graph depicting individual patient (dots) and median (thick black line) TNF $\alpha$  release.

<https://doi.org/10.1371/journal.pone.0226564.g002>

combination therapy. This was observed as a reduction in the level of TNF $\alpha$  released from the biopsies into the culture media (Fig 2). Different levels of response were however observed between patients and ranged from modest to a marked reduction in TNF $\alpha$  in the supernatant.

Fluticasone alone or in combination with roflumilast generated the greatest inhibition of TNF $\alpha$  release. When the effects of monotherapy and combined therapy were compared, there was no difference in the mean reduction in TNF $\alpha$  levels; however, combined therapy may have resulted in a bimodal pattern of drug response across the patient sample group (Fig 3). Based on this density plot, patients were dichotomised into "High responders" and "Low responders" based on the average values. Twelve patient samples were categorised as being high responders and thirteen as low responders to the roflumilast plus fluticasone treatment.

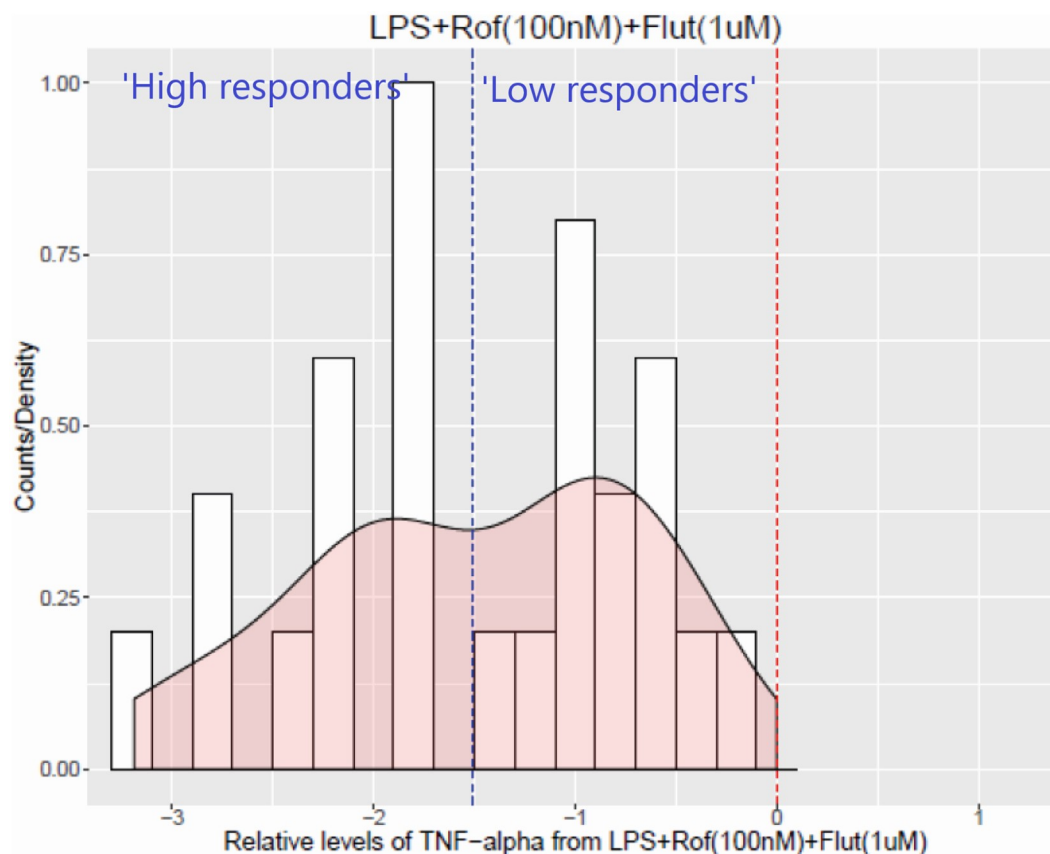
Principal component analysis [19] was conducted to explore the relationship between the many variables (Fig 4).

Association analysis of patient demographic parameters and response to roflumilast plus fluticasone showed that the response was not influenced by any of the patient demographic factors such as gender or age. Treatment response was noted to be significantly associated with the first principal component; this indicates that response to roflumilast plus fluticasone is the primary trend in the data.

A strong association was observed between ethnicity and supplier region, this is however believed to be the result of one sample that was acquired from a geographical region distinct from all other regions. The ethnicity of this patient was also not replicated in any other sample.

Classes of chronic medication appear to be strongly related to each other; this is not surprising as the standard of care treatment for COPD includes combinations of the classes of drugs identified. Chronic medication appeared not to influence patient response to roflumilast plus fluticasone in the organoculture assay and is therefore not thought to be responsible for the variation in response between patients.





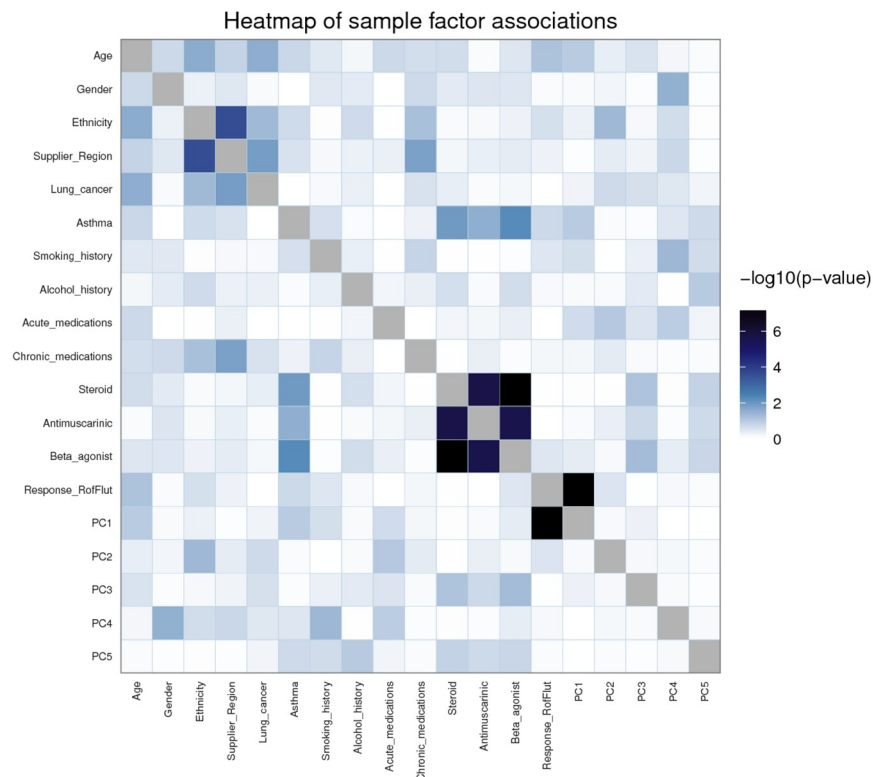
**Fig 3. Visualisation of patient-to-patient changes in the relative levels of TNF $\alpha$  after combination treatment of roflumilast (100 nM) and fluticasone (1  $\mu$ M).** The density plot shows histogram bin counts (number of times a value falls within a given bin) as white bars as well as a smooth density in pink of the log<sub>2</sub> ratios of TNF $\alpha$  release from biopsies treated with roflumilast plus fluticasone, relative to those treated with DMSO (which also includes LPS and the vehicle control), across all 25 patients. The average level in the treated biopsies is denoted by a blue dashed vertical line and the red dashed line denotes zero as this is the average level in control biopsies.

<https://doi.org/10.1371/journal.pone.0226564.g003>

### Exome sequence analysis

Preliminary analysis showed that all samples were of good quality, with between 38 and 57 million reads; this resulted in 36,702 to 38,065 SNPs being identified per patient sample. Merging and filtering of VCF files for high quality SNPs resulted in 101,557 SNPs being retained for the exome wide association analysis. Hierarchical clustering and principal component analysis identified two patient samples as outliers. One of the outlying samples is described above and is thought to have resulted in a slight association with ethnicity and supplier region. There is no explanation for the second outlying sample, however as the two samples did not show any quality-related discrepancies both samples were included in downstream analysis.

Fisher's exact test, performed in the PLINK toolkit, showed that no genotypes corresponding with the identified SNPs were significantly associated with the organoculture response. However, to allow a very tentative interpretation of the results, and taking into account the low number of patients studied, an uncorrected p-value of <0.001 was chosen. With this approach a total of 30 SNPs, corresponding to 23 genes, were found to correlate with the level of TNF $\alpha$  release upon treatment with roflumilast plus fluticasone. A number of these genes have reported associations with COPD or other pulmonary diseases and include; CYP2E1 [20], HEY1 [21], SMAD3 [22], BARD1 [23] and FOXP1 [24].



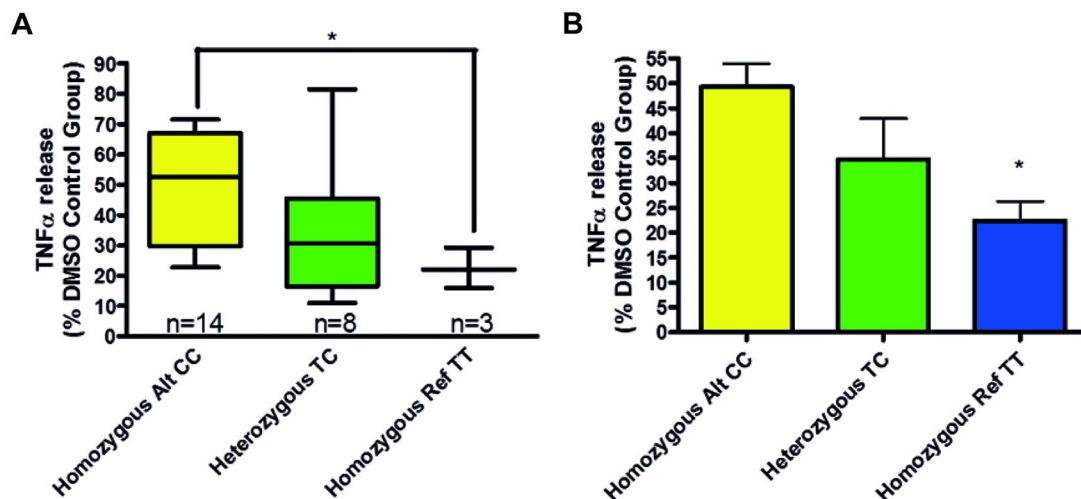
**Fig 4. Heat map showing the results of patient demographic correlation with TNF $\alpha$  release from stimulated human lung parenchyma biopsies following roflumilast and fluticasone co-treatment.** Each parameter is assessed in relation to each other, the principal components (PC) driving variation in the data and to the organoculture assay response. Each area within the heatmap denote a p-value of association between pairs of variables from statistical tests. The statistical tests utilised depends on the property of the factors: for an association between two categorical factors, a chi-squared test was used. For an association between a categorical and a continuous factor, ANOVA was used. For an association between two continuous factors, a Spearman correlation test was used. In all cases, the resulting p-value was transformed as  $-\log_{10}(p)$  before being visualised in the confounding factors heatmap.

<https://doi.org/10.1371/journal.pone.0226564.g004>

CYP2E1 is an inducible drug metabolising enzyme expressed in human lung tissue and has been implicated in pathological oxidative stress [20,25]. Expression of CYP2E1 SNPs including rs3737034 and rs2249695 were shown to correlate with patient organoculture response. The significance level was however borderline as determined in the bioinformatic analysis ( $p = 0.008/0.01$ ).

Our findings suggest that genetic variation in the cytochrome p450 enzyme (CYP2E1) gene, namely SNP (rs2249695), may partly explain the observed variation in drug response. Biopsies from patients who had at least one copy of the reference allele for this SNP generally responded better to roflumilast and fluticasone co-treatment. Mean TNF $\alpha$  release (Fig 5) was inhibited by 77.6% (homozygous reference genotype (TT)) and by 50.74% (homozygous alternative genotype (CC)). Levels of inhibition between these two genotypes were found to be significantly different with a p value of 0.02 (unpaired, two-tailed t-test). The homozygous reference haplotype has been associated with low CYP2E1 expression [20].

Genetic variation in mothers against decapentaplegic homolog 3 (SMAD3) gene was also found to relate to patient organoculture response. Mean TNF $\alpha$  release (Fig 6) was inhibited by 66% (homozygous alternative genotype (GG)) and by 39% (heterozygous genotype). Levels of inhibition between these two genotypes were found to be significantly different with a p value of 0.0054 (unpaired, two-tailed t-test). Only two patient samples were found to have the

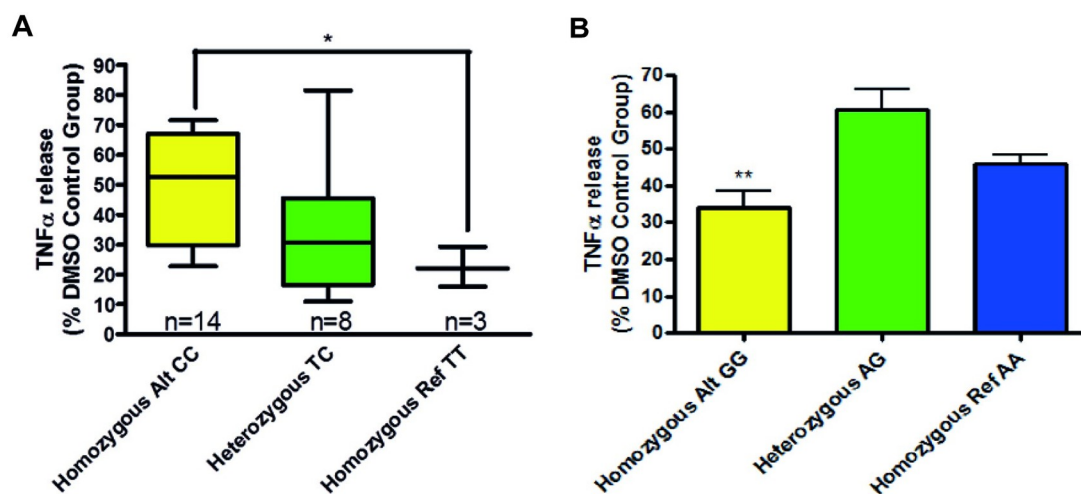


**Fig 5. Graphs showing the relationship between CYP2E1 SNP rs2249695 genotype and TNF $\alpha$  release from stimulated human lung parenchyma biopsies following roflumilast and fluticasone co-treatment.** Data is displayed as a percentage of the corresponding DMSO control group. Asterisks indicate significant differences ( $p < 0.05$ , for one,  $p < 0.01$  for two and  $p < 0.001$  for three). A: Box and whiskers graph depicting TNF $\alpha$  release. The 25th and 75th percentiles of each group are represented by the box with the minimum and maximum values represented by bars, the line within each box denotes the median value. B: Bar graph depicting mean + SEM TNF $\alpha$  release.

<https://doi.org/10.1371/journal.pone.0226564.g005>

homozygous reference haplotype (AA) and mean TNF $\alpha$  release was inhibited by 54% in this group of patient samples. This level of inhibition was not significantly different to the homozygous alternative genotype or the heterozygous genotype.

The GWAS catalogue contains 624 SNPs identified in the exome sequence analysis, 4 of these SNPs are annotated in the catalogue as being associated with COPD; 6 have been associated with asthma and 4 are related to other pulmonary conditions. It was however found that



**Fig 6. Graphs showing the relationship between SMAD3 SNP rs1065080 genotype and TNF $\alpha$  release from stimulated human lung parenchyma biopsies following roflumilast and fluticasone co-treatment.** Data is displayed as a percentage of the corresponding DMSO control group. Asterisks indicate significant differences ( $p < 0.05$ , for one,  $p < 0.01$  for two and  $p < 0.001$  for three). A: Box and whiskers graph depicting TNF $\alpha$  release. The 25th and 75th percentiles of each group are represented by the box with the minimum and maximum values represented by bars, the line within each box denotes the median value. B: Bar graph depicting mean + SEM TNF $\alpha$  release.

<https://doi.org/10.1371/journal.pone.0226564.g006>

no SNPs annotated in the catalogue correlated to roflumilast plus fluticasone response in this study.

### miRNA analysis

RNA quality control analysis showed that isolated RNA was of high purity, 260/280 ratios ranged from 1.8 to 2.0 and RNA integrity scores ranged from 5.8 to 7.8.

All patient samples, except one, passed Agilent miRNA array quality control analysis with a rating of good to excellent. The remaining sample was flagged for evaluation and removed from subsequent bioinformatic analysis.

Statistical analysis showed that there were no specific differences in miRNA expression detected in patients that responded well in the organoculture assay compared with patients that did not respond. This analysis was performed using a p-value that had been adjusted for multiple statistical testing. For the purposes of this exemplar study, a relaxed p-value (uncorrected  $p < 0.05$ ) was subsequently applied. At this threshold, 181 miRNAs, mapping to 636 genes, were found to be differentially expressed in COPD patient samples that were high responders to roflumilast plus fluticasone treatment compared with samples that showed a poor response. 86 miRNAs were found to be upregulated, correlating with 47 KEGG pathways that reached statistical significance. This Enrichment analysis highlighted KEGG pathways associated with TGF- $\beta$  signalling, synaptic function and fatty acid metabolism.

95 miRNAs were found to be down regulated correlating with 4 KEGG pathways that reached statistical significance. This Enrichment analysis highlighted KEGG pathways associated with long-term depression and serotonergic and GABAergic synaptic function.

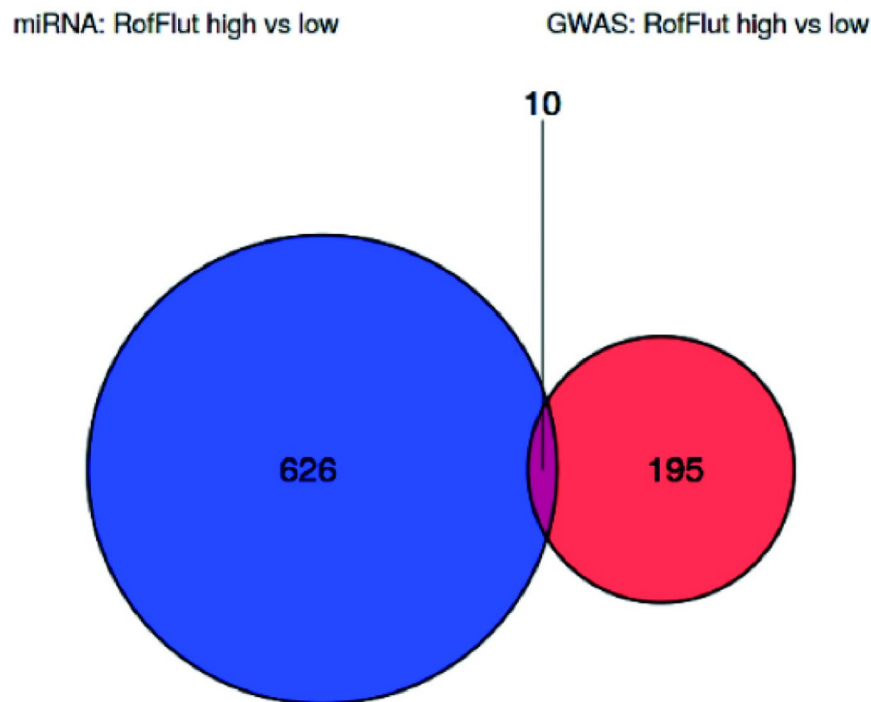
1,610 GO terms were significantly associated with up-regulated miRNAs and found to be significantly enriched for pathways associated with cell ageing, specifically telomerase activity. Pathways involved in synaptic activity and T cell differentiation were also found to be upregulated.

310 GO terms were significantly associated with down-regulated miRNAs and found to be significantly enriched for pathways associated with B cell receptor activity and TGF- $\beta$  production.

As discussed, bioinformatic analysis identified 30 SNPs corresponding to 23 genes ( $p < 0.001$ ) and 181 miRNAs (mapping to 636 genes,  $p < 0.05$ ) as being related to organoculture response. With further relaxation of the exome analysis p value to 0.01, congruence analysis found that a total of 10 genes overlapped between the exome sequence and miRNA expression data (Fig 7). This overlap is higher than would be expected by chance. Overlapping genes are NTN4, IGF1R, SMAD3, EGFR, MCL1, FBN1, FGA, APP, MYO10 and IRAK3. Six overlapping genes were subject to upregulation (SMAD3, EGFR, MCL1, FBN1, FGA & APP) however the remaining 4 overlapping genes did not agree with respect to overlap direction. Absolute minor allele frequencies from the exome sequence analysis were used as a surrogate for fold changes in the SNP data. No strong correlations were found between absolute minor allele frequencies and miRNA log fold-changes. KEGG and GO enrichment analysis of the overlapping genes did not identify any common pathways or processes.

### Discussion

This study aimed to demonstrate the potential of research that combines preclinical functional characterisation of drug efficacy and inter-patient variation in drug responses, with state-of-the-art genomics and bioinformatics, as a new way to model precision medicine strategies at the early stages of drug development.



**Fig 7. Venn diagram illustrating the overlap between genes that map to SNPs and miRNAs that are associated with TNF $\alpha$  release from stimulated human lung parenchyma biopsies following roflumilast and fluticasone co-treatment.**

<https://doi.org/10.1371/journal.pone.0226564.g007>

COPD is a highly complex condition with many clinical phenotypes. As an exemplar project, the number of patients was relatively low and findings are therefore tentative, however, this study was also designed to explore the potential for such projects during non-clinical drug development, where budgets are limited and projects exploring hundreds of patients may be too costly.

Nonetheless, clear variations in drug effectiveness were observed between patients and our preliminary experimental findings suggest that genetic polymorphisms in COPD patients may be linked to variation in response to the combination anti-inflammatory treatment, roflumilast plus fluticasone. A haplotype associated with low CYP2E1 expression was detected within the cohort of samples that responded well to treatment. It is therefore possible that CYP2E1 expression influences response to treatment.

CYP2E1 induces production of reactive oxygen species [20,25] that may in turn inhibit reductions of TNF $\alpha$  release by various treatments. All 3 patients in the homozygous reference haplotype group were high responders to roflumilast plus fluticasone, 5 of 8 patients in the heterozygous reference haplotype group were high responders whereas 10 of 14 patients in the homozygous alternative haplotype group were low responders (Fig 5).

TGF- $\beta$  and the SMAD signalling pathway have been implicated in the pathology of COPD [26,27] and lung adenocarcinoma [28,29]. Our results show that genetic variation in the SMAD gene (rs1065080) may influence response to fluticasone plus roflumilast. Patients that were deemed to be high responders to roflumilast plus fluticasone exclusively displayed the homozygous alternative genotype (GG), whereas only 5 of 13 patients in the poor response group displayed this genotype.



Roflumilast has been reported to inhibit TGF- $\beta$  driven increases in reactive oxygen species and phosphorylation of SMAD3 by inhibiting TGF- $\beta$  release [30]. If the genetic variation in SMAD3 and miRNA expression profile reported in this study alters the functioning of the pathway then this may help to explain variation in the observed organoculture response. It was however noted that no common KEGG or GO pathways were found in the bioinformatic congruence analysis.

The AnalytiXagility platform used by partners to share and interrogate the data could become a powerful resource to both academic researchers and the pharmaceutical industry.

Aridhia's digital research platform has the potential to link the data generated in this study with available tissue, DNA and RNA for further research. A platform of this design also offers the capacity to add patients, analyses and clinical information in real time, thereby tracking patient outcome and allowing continual remodelling of the data in a secure, version controlled manner.

With ethical approval, it could be possible for researchers in the pharmaceutical industry to mine for genetic signatures or other parameters within a target disease area, for the purposes of patient selection and clinical trial support or for identifying the most appropriate preclinical model.

The authors acknowledge that while a very high volume of functional and genomics data was generated, the total number of patients was low for a genomics study and that the statistical thresholds varied to allow some associations to be made between the data sets. For this reason, the scientific conclusions made from the study remain tentative. The authors do however feel that this project serves to demonstrate well the potential to explore patient stratification strategies at a much earlier stage by combining fresh tissue pharmacology, clinical metadata and genomics.

## Author Contributions

**Conceptualization:** David C. Bunton.

**Data curation:** Jane Hair, Max Bylesjo.

**Formal analysis:** Graeme Macluskie, Michael Finch, Colin N. A. Palmer, Max Bylesjo, Carolyn Low, David Mallinson, Elaine M. Gourlay, Linda Cheyne.

**Methodology:** Max Bylesjo.

**Project administration:** Hannah Child.

**Resources:** Jane Hair, Pamela Brankin.

**Software:** Max Bylesjo.

**Supervision:** Graeme Macluskie, Sarah Lynagh, Marian McNeil, Hannah Child, David C. Bunton.

**Visualization:** Colin N. A. Palmer.

**Writing – original draft:** Karen Cowan, David C. Bunton.

**Writing – review & editing:** Karen Cowan, Graeme Macluskie, David C. Bunton.

## References

1. Agyeman AA, Ofori-Asenso R. Perspective: Does personalized medicine hold the future for medicine? J Pharm Bioallied Sci. 7: 239–44. <https://doi.org/10.4103/0975-7406.160040> PMID: 26229361

2. Chen Y, Guzauskas GF, Gu C, Wang BCM, Furnback WE, Xie G, et al. Precision Health Economics and Outcomes Research to Support Precision Medicine: Big Data Meets Patient Heterogeneity on the Road to Value. *J Pers Med*. 2016; 6. <https://doi.org/10.3390/jpm6040020> PMID: 27827859
3. Paul SM, Mytelka DS, Dunwiddie CT, Persinger CC, Munos BH, Lindborg SR, et al. How to improve RD productivity: The pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*. 2010. pp. 203–214. <https://doi.org/10.1038/nrd3078> PMID: 20168317
4. Schork NJ. Personalized medicine: Time for one-person trials. *Nature*. Nature Publishing Group; 2015. pp. 609–611. <https://doi.org/10.1038/520609a> PMID: 25925459
5. Trusheim MR, Burgess B, Hu SX, Long T, Averbuch SD, Flynn AA, et al. Quantifying factors for the success of stratified medicine. *Nat Rev Drug Discov*. 2011; 10: 817–33. <https://doi.org/10.1038/nrd3557> PMID: 22037040
6. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. *Nat Rev Drug Discov*. 2014; 13: 419–31. <https://doi.org/10.1038/nrd4309> PMID: 24833294
7. Biankin AV, Piantadosi S, Hollingsworth SJ. Patient-centric trials for therapeutic development in precision oncology. *Nature*. 2015; 526: 361–70. <https://doi.org/10.1038/nature15819> PMID: 26469047
8. Found in translation—European Pharmaceutical Review. [cited 30 Oct 2019]. <https://www.europeanpharmaceuticalreview.com/article/16219/found-in-translation/>
9. Asghar W, El Assal R, Shafiee H, Pitteri S, Paulmurugan R, Demirci U. Engineering cancer microenvironments for in vitro 3-D tumor models. *Mater Today (Kidlington)*. 2015; 18: 539–553. <https://doi.org/10.1016/j.mattod.2015.05.002> PMID: 28458612
10. Zanon M, Piccinini F, Arienti C, Zamagni A, Santi S, Polico R, et al. 3D tumor spheroid models for in vitro therapeutic screening: A systematic approach to enhance the biological relevance of data obtained. *Sci Rep*. 2016; 6. <https://doi.org/10.1038/srep19103> PMID: 26752500
11. Mandrekar SJ, Sargent DJ. All-Comers versus Enrichment Design Strategy in Phase II Trials. *J Thorac Oncol*. 2011.
12. Cilli A, Bal H, Gunen H. Efficacy and safety profile of roflumilast in a real-world experience. *J Thorac Dis*. 2019; 11: 1100–1105. <https://doi.org/10.21037/jtd.2019.04.49> PMID: 31179051
13. Rennard SI, Tashkin DP, McElhattan J, Goldman M, Ramachandran S, Martin UJ, et al. Efficacy and tolerability of budesonide/formoterol in one hydrofluoroalkane pressurized metered-dose inhaler in patients with chronic obstructive pulmonary disease: results from a 1-year randomized controlled clinical trial. *Drugs*. 2009; 69: 549–65. <https://doi.org/10.2165/00003495-200969050-00004> PMID: 19368417
14. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004; 5: R80. <https://doi.org/10.1186/gb-2004-5-10-r80> PMID: 15461798
15. Zheng X. Statistical Prediction of HLA Alleles and Relatedness Analysis in Genome-Wide Association Studies. 2013.
16. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81: 559–575. <https://doi.org/10.1086/519795> PMID: 17701901
17. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007; 8: 118–127. <https://doi.org/10.1093/biostatistics/kxj037> PMID: 16632515
18. Chou C-H, Chang N-W, Shrestha S, Hsu S-D, Lin Y-L, Lee W-H, et al. miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res*. 2016; 44: D239–47. <https://doi.org/10.1093/nar/gkv1258> PMID: 26590260
19. Edward Jackson J, York Chichester Brisbane Toronto N. A User's Guide To Principal Components.
20. Arif E, Vibhuti A, Alam P, Deepak D, Singh B, Athar M, et al. Association of CYP2E1 and NAT2 gene polymorphisms with chronic obstructive pulmonary disease. *Clin Chim Acta*. 2007; 382: 37–42. <https://doi.org/10.1016/j.cca.2007.03.013> PMID: 17442289
21. Tilley AE, Harvey B-G, Heguy A, Hackett NR, Wang R, O'Connor TP, et al. Down-regulation of the notch pathway in human airway epithelium in association with smoking and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2009; 179: 457–66. <https://doi.org/10.1164/rccm.200705-795OC> PMID: 19106307
22. Faura Tellez G, Vandepoele K, Brouwer U, Koning H, Elderman RM, Hackett T-L, et al. Protocadherin-1 binds to SMAD3 and suppresses TGF- $\beta$ 1-induced gene transcription. *Am J Physiol Lung Cell Mol Physiol*. 2015; 309: L725–35. <https://doi.org/10.1152/ajplung.00346.2014> PMID: 26209277

23. André PA, Prêle CM, Vierkotten S, Carnesecchi S, Donati Y, Chambers RC, et al. BARD1 mediates TGF- $\beta$  signaling in pulmonary fibrosis. *Respir Res*. 2015; 16. <https://doi.org/10.1186/s12931-015-0278-3> PMID: 26415510
24. Chokas AL, Trivedi CM, Lu MM, Tucker PW, Li S, Epstein JA, et al. Foxp1/2/4-NuRD interactions regulate gene expression and epithelial injury response in the lung via regulation of interleukin-6. *J Biol Chem*. 2010; 285: 13304–13. <https://doi.org/10.1074/jbc.M109.088468> PMID: 20185820
25. Linhart K, Bartsch H, Seitz HK. The role of reactive oxygen species (ROS) and cytochrome P-450 2E1 in the generation of carcinogenic etheno-DNA adducts. *Redox Biol*. 2014; 3: 56–62. <https://doi.org/10.1016/j.redox.2014.08.009> PMID: 25462066
26. Yang T, Ying B, Song X, Zhang S, Fan H, Xu D, et al. Single-nucleotide polymorphisms in SMAD3 are associated with chronic obstructive pulmonary disease. *Exp Biol Med (Maywood)*. 2010; 235: 599–605. <https://doi.org/10.1258/ebm.2010.009268> PMID: 20463300
27. Ezzie ME, Crawford M, Cho J-H, Orellana R, Zhang S, Gelinas R, et al. Gene expression networks in COPD: microRNA and mRNA regulation. *Thorax*. 2012; 67: 122–31. <https://doi.org/10.1136/thoraxjnl-2011-200089> PMID: 21940491
28. Yang H, Wang L, Zhao J, Chen Y, Lei Z, Liu X, et al. TGF- $\beta$ -activated SMAD3/4 complex transcriptionally upregulates N-cadherin expression in non-small cell lung cancer. *Lung Cancer*. 2015; 87: 249–57. <https://doi.org/10.1016/j.lungcan.2014.12.015> PMID: 25595426
29. Yu J-R, Tai Y, Jin Y, Hammell MC, Wilkinson JE, Roe J-S, et al. TGF- $\beta$ /Smad signaling through DOCK4 facilitates lung adenocarcinoma metastasis. *Genes Dev*. 2015; 29: 250–61. <https://doi.org/10.1101/gad.248963.114> PMID: 25644601
30. Milara J, Peiró T, Serrano A, Guijarro R, Zaragoza C, Tenor H, et al. Roflumilast N-oxide inhibits bronchial epithelial to mesenchymal transition induced by cigarette smoke in smokers with COPD. *Pulm Pharmacol Ther*. 2014; 28: 138–48. <https://doi.org/10.1016/j.pupt.2014.02.001> PMID: 24525294